



OIPE

#2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/838,044

DATE: 01/25/2002

TIME: 16:50:53

Input Set : N:\Crf3\RULE60\09838044.raw

Output Set: N:\CRF3\01252002\I838044.raw

1 <110> APPLICANT: Kaser, Matthew R.
 2 Azimzai, Yalda
 3 Yue, Henry
 4 <120> TITLE OF INVENTION: POLYCYCLIC AROMATIC HYDROCARBON INDUCED MOLECULES
 5 <130> FILE REFERENCE: PB-0011 US
 6 <140> CURRENT APPLICATION NUMBER: 09/838,044
 7 <141> CURRENT FILING DATE: 2001-04-18
 9 <150> PRIOR APPLICATION NUMBER: 09/386,493
 10 <151> PRIOR FILING DATE: 1999-08-30
 13 <160> NUMBER OF SEQ ID NOS: 14
 14 <170> SOFTWARE: PERL Program
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 2439
 18 <212> TYPE: DNA
 19 <213> ORGANISM: Homo sapiens
 20 <220> FEATURE:
 W--> 21 <221> NAME/KEY:
 22 <223> OTHER INFORMATION: 1851405
 23 <300> PUBLICATION INFORMATION:
 24 <400> SEQUENCE: 1

```

25      cttcagttcc tcgaagagat tcactttcaa aaacatcaac tccttataac aaatcaaaca 60
26      aagcagcaag ccaacaaggg accccatggg aaacacttgt cgtgtttgct atcaacttga 120
27      agcaattaaa cggtcacaatg aatatgagta atgtaatggg aaatacaact tggacaacta 180
28      gtggtttgaa gagccagggc cgtctgtcag taggaagtaa tcgtgatcga gagatcagca 240
29      tgtctgttggt tctgggaaga tcacaattag attctaaagg aggagtagtt ggagggacca 300
30      tagatgtcaa tgctttggag atgggttgctc atatttctga acatccaaat cagcaaccga 360
31      gtcacaaaat tcagattact atgggttcta ctgaagctcg tgttgattac atgggctcaa 420
32      gtatcctcat gggcatcttc agtaatgctg atcttaagct tcaggatgaa tggaaagtaa 480
33      acttgtataa tacattggat tcaagcataa ctgataaaag tgagattttc gtccatggag 540
34      atttgaagtg ggataatttc caagtaatga tatcaaggctc aaccacacca gatctgataa 600
35      aaataggaat gaagctccag gaatttttca cacaacaatt tgataaccagc aaacgagctc 660
36      tgtctacctg gggaccagtt ccttaccttc cgccaaagac aatgactagc aacctagaaa 720
37      aaagttcaca agaacaatta ctgtatgcag cacatcatcg aactggcctt ggagtattga 780
38      aggtggtatc aggatgccac atatccttat ttcagattcc attaccagaa gatggaatgc 840
39      aatttgaggg atcaatgagc ttacatggaa atcatatgac actggcatgt tttcatggtc 900
40      caaattttcg ttcaaaatct tgggcocttt ttcatctaga agaaccaaat attgcttttt 960
41      ggactgaagc tcagaaaatc tgggaagatg gctccagtga tcattctaca tatattgtac 1020
42      aaacactaga ttttcacctg ggtcataata ctatggttac caaacatgtt ggtgcttttg 1080
43      aaagtcctat ggcaacaata accaagataa caaggcgtcg ccatgaaaat ccaccccatg 1140
44      gagtagcaag tgtgaaagaa tggttcaatt atgttacagc tacaaggaat gaagagctaa 1200
45      atctgcttcg taatgttgat gctaacaaca ctgagaatag cactactgtg aagaattcta 1260
46      gtttggttgag tggattcaga ggaggttcta gctacaacca tgaaacagag actatctttg 1320
47      cattaccaag gatgcagctt gactttaaat ccattcatgt tcaagaacca caggagcctt 1380

```

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/838,044

DATE: 01/25/2002

TIME: 16:50:53

Input Set : N:\Crif3\RULE60\09838044.raw

Output Set: N:\CRF3\01252002\I838044.raw

```

48   cattacagga tgccagcctg aagccaaaag tagaatgtag tgtggtgaca gagttcactg 1440
49   accacatttg tgtgactatg gatgctgagc tcatcatggt tcttcatgat ttagtatcag 1500
50   cttatcttaa agaaaaagaa aaagccatct tccacacctg gattttatct actcgaccag 1560
51   gacaaaaaag tccaattatt atacatgacg acaattcctc tgataaagat agagaagata 1620
52   gcatcactta tactactgtg gactggagag attttatgtg caatacatgg cacctagaac 1680
53   ctactcttag attaatctct tggactggaa gaaagattga tccagtaggt gttgattata 1740
54   ttcttcaaaa attgggcttt catcatgcta ggactactat tcctaaatgg cttcaaagag 1800
55   gagtcatgga tccactggac aagggttctgt cagttcttat caaaaagctc ggtactgcac 1860
56   tacaggatga aaaggaaaag aaaggcaaag acaaagaaga acactaaaaa agtaatttga 1920
57   tctgtgaaca aattatgatt gtgtctgttt tattacactg gagtggtttt ttagtataat 1980
58   aatttgaaat ataactttta aataattcta aatttggtgc tataattaaa agtttgtaag 2040
59   ttaacctgtt ctagttccat cattctgtgt acagtgaagt attgcatgat aatgtaaatt 2100
60   ttgtgaaaaa ctagattaaa atatatact gcttgttatg gtttataatt atataatgtg 2160
61   caatacaatt cctgcatctt taaaatgtct gcagaataac tgtgaatttt ttgtttattg 2220
62   gattggccgt aactttttag aaaaaatctt gttgatgata atgtgatttt ggggaggtca 2280
63   ttaattgctt tttctttttt aaatgtagac ttatataaat acctgtttgt atatagcttg 2340
64   agtaattgtg atatgattgt ataccactaa aatattgtta actattataa taaagtcaca 2400
65   gtaatggttt aaaaaagaaa aaagaaaaaa aaaaaaaaaa 2439

```

67 <210> SEQ ID NO: 2

68 <211> LENGTH: 529

69 <212> TYPE: DNA

70 <213> ORGANISM: Homo sapiens

71 <220> FEATURE:

W--> 72 <221> NAME/KEY:

73 <223> OTHER INFORMATION: 1991226

74 <300> PUBLICATION INFORMATION:

75 <400> SEQUENCE: 2

```

76   ggacgctggc aaagcggggc agaccctgaa gactcactgc tcagcccagc gccagatgt 60
77   ctgcaggtgg ctgagcccct tcatcctctc ctgctgcgtg tacttctgcc tctggattcc 120
78   cgaggaccag ctgtcctggt tcgtgcctct ggtcaagtgc ctgcctgtcc tctgcctggc 180
79   tgggttcctg tgggtcatgt ccccaagcgg gggctacacc cagctcctcc agggagccct 240
80   tgtgtgctcg gctgtggggg acgcttgccct catctggccg gcagccttcg tccctggcat 300
81   ggcgccttt gccaccgccc acctcctcta cgtctgggcc ttcggcttct ctcccctgca 360
82   gcccgccctg ctgctgctca tcatcctggc ccctggcccc tactcagctt gtgctcagca 420
83   gctcgagccg gataggtctg cgggtggagct atggctatct gatgcatctg tgcgggctgc 480
84   cacgtcgatg cgctgggcgg tgtctcacgt tctatggtgt gctggaact 529

```

86 <210> SEQ ID NO: 3

87 <211> LENGTH: 2429

88 <212> TYPE: DNA

89 <213> ORGANISM: Homo sapiens

90 <220> FEATURE:

91 <221> NAME/KEY: unsure

92 <222> LOCATION: 2404, 2413, 2423, 2425, 2426

93 <223> OTHER INFORMATION: a or g or c or t, unknown, or other

94 <220> FEATURE:

W--> 95 <221> NAME/KEY:

96 <223> OTHER INFORMATION: 253053

97 <300> PUBLICATION INFORMATION:

98 <400> SEQUENCE: 3

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/838,044

DATE: 01/25/2002

TIME: 16:50:53

Input Set : N:\Crif3\RULE60\09838044.raw

Output Set: N:\CRF3\01252002\I838044.raw

```

99      ttatgtgttt aacctataaa tattgggggtc ttctgtctaa actgggggtca ctgttgcatg 60
100     gaacattgtt cttaatagtt gagaattgct tttttgaaaa ttttcatgaa ggaatttttg 120
101     taatgacttt gcttgacagt tttttggggt gttttgagaa agtggcatgg aaacatgcag 180
102     tagttaatga gtttctcttg gtactgaaca ctattagaat atcattagtg atattttttc 240
103     tcttttagagc atttttaaatg caactagccc ctatatttta atgtaagagt tactctgcaa 300
104     tctaagcaaa gcaccaaca atggtaaagt ttttttaaaa atgcagaact aagatttttg 360
105     actctaaaga gagaaaatta caaggggtgt gccttatagc aaaccttggt gacaatcctt 420
106     catgtgagca aagtgttgat cttaatatgt gttgtctgtg gtgtgctttt ttgtactgta 480
107     aaaatatgtg gttcatgtct aactctgctg ttttattgtg gttgtgggtc aagtttttaa 540
108     tgtttaaagt tgatgctgtt ttcagaagag ctttttacta atttatttgt cagtgttccc 600
109     tatttgttac ttaaccatga tcttcagat tttttggagt attcttttct aaccttaacc 660
110     ctgccaacc ttgatccatt ttgacatttg ttatgcacta tttttatata tctgtgagag 720
111     atttttccaa cagtcageta ttttatggca cactttttt gactgatgac atctccttg 780
112     ctatacctca atttttggaa tttagagaag aaatcagtag ttttgcaatg ttaattattt 840
113     agatattcaa ttttcagat ttttaactt tattttcata atttctgctt aatgttttaa 900
114     attgaagagc cttttcatgt attaaataat gaacacaaat tatataatta aaataattgg 960
115     agatgttgaa aatcattttt cttctttaa cagaaataaa tatttggaaat gaaggggaat 1020
116     gtactagaac accctttttg ccacgggtaa aaataacaga aatgtatggt ttgttttacc 1080
117     ttcatttctg tacaagtaaa gcttattagt ctaatgtttt gttcctttcc cacctcaccc 1140
118     ctacctottt tgttttgttt tgttttggcc ctttatgtac tacattctta ttttctaact 1200
119     tttaaacact gtattggagg tttttttttt taatttacag atcatattta ttttactatt 1260
120     tttgtagaaa attattaatt ttgattgtat ttttgtattt taaaagcttc ttcaottgtg 1320
121     ttccctaaat attcatattg ctgccccaaa gtatgactgt ggaggaaaaa aaaatacttt 1380
122     aaaaatccac actttttgtt aagaaggaaa catttagcat ttatatattt gtgtatggaa 1440
123     aacacttgat attttatccc tgttgcatct ggtgcacag agcctctcct caaagatgct 1500
124     acaaaacttg aatataacac attttggaag gctgactaac ctcgattctg tgttgtgatg 1560
125     tgcaatactg tttctaagt ttgtataaaa aaaaacagt taaacctttt taatgcaaat 1620
126     ttattttttt cattgcatat ttgcagatt ttatccacag tgtcattttt tactgtcaga 1680
127     aaagataccc cttttgtcat tgcaactatt ttttaaatcc agaaatcttt gtaactgatg 1740
128     taaatgattg tagttatttt gggatagtgt tttggctaac aaaagggaga gacttttttc 1800
129     atggcatatt tctattttgt ttttttgggt tttattttat tttaatagtt agtaaaatac 1860
130     ttgggaataa tttttgcata ttcttgttca ttaatattat tttgtatttt tatgtggaaa 1920
131     tatataattt tatgacacta attgctaaag tttattttat gttgaattat ttttggagct 1980
132     gaaatctttg taatattaaa gcaactagtt tctaattccc agtttctgta tagaatcgca 2040
133     caagtgggtt atggagtgtt tggattgtaa ttataaatgg ttctttgata tgcaaatata 2100
134     tattttcagt tgattttatt ttatattcct aatgggggtg taaagccgtt ttttattttt 2160
135     ttctaaataa aaagagaacc catgctttta tggacactag gtaaacacct tcagcttaaa 2220
136     tttttcgtta aatattttag tttattttat tgttatcttc caggtgtcta aatctccagt 2280
137     ctgtctgttg tactggtaat ttaactctgt aatggaatag tttgctgcca actatttata 2340
138     ttaagtaatt tttaaatatt tgtaatatgt ttgactgact aataaactat taagttattg 2400
W--> 139     gaangaaaaa canaaaaaaa aanannaac 2429
141 <210> SEQ ID NO: 4
142 <211> LENGTH: 944
143 <212> TYPE: DNA
144 <213> ORGANISM: Homo sapiens
145 <220> FEATURE:
W--> 146 <221> NAME/KEY:
147 <223> OTHER INFORMATION: 2009569
148 <300> PUBLICATION INFORMATION:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/838,044

DATE: 01/25/2002

TIME: 16:50:53

Input Set : N:\Crf3\RULE60\09838044.raw

Output Set: N:\CRF3\01252002\I838044.raw

149 <400> SEQUENCE: 4
 150 gcttgaaggc gggaaaatac ggaatggagc gacagggagt ggacgtgcc catgtgaaat 60
 151 gcaaagacca ggaaccgcag cccttggggg agagcaagga gcatccgcg tgggaagaga 120
 152 actgcgagga ggaagctggt ggagggccag ctagtgccag ttgccagctg acggtcctgg 180
 153 aagggaaagtc gggactctac ttctcctctc tggactcaag cattgacatc ctgcagaaga 240
 154 gagcccagga gctgatcgaa aacatcaaca agagccggca aaaggaccat gcaactcatga 300
 155 ccaacttcag gaacagcctg aagaccaagg ttctcgatct gacagagaaa ttagaggaga 360
 156 ggatctatca gatttataat gaccacaaca agatcatcca ggaaaagctc caagagttca 420
 157 ccagaaaaat ggcaaagatc agccatttgg agacagagct caaacaagtc tgccacagcg 480
 158 tggagactgt gtacaaagac ctgtgtctcc agcctgagca gaggcctaaga ctcatagggg 540
 159 ggccagacca ctctagggga aagtccccac cacgtcccgg caactcacag ccccagacg 600
 160 tgttcgtttc ttctgtggct gaaactactt ctacggccac tgcttcagaa gtacagacca 660
 161 acagagatgg tgaatgctga cagctgccgg gagactcacg ccttagtgac agtctccagg 720
 162 agaagactgt gaggccacca tttgggccac actgagaaat tgtttttcat ggttctataa 780
 163 tgcattcttg cagaaaaaaa caaaaaccca aagctccttg tgcgaactc caaaaatgta 840
 164 gcaagtccag cctctccat aggccaggc ttctgtctcc ccacccttgg caagttctcc 900
 165 ccacccccag cccacagtt tattaatgt ttgattttca aaaa 944

167 <210> SEQ ID NO: 5

168 <211> LENGTH: 404

169 <212> TYPE: DNA

170 <213> ORGANISM: Homo sapiens

171 <220> FEATURE:

W--> 172 <221> NAME/KEY:

173 <223> OTHER INFORMATION: 1642580

174 <300> PUBLICATION INFORMATION:

175 <400> SEQUENCE: 5

176 ctcaagaccc agcagtggga cagccagaca gacggcacga tggcactgag ctcccagatc 60
 177 tgggcccgtt gctcctgct cctcctcctc ctgcagacc tgaccagtgg ctctgttttc 120
 178 ccacaacaga cgggacaact tgcagagctg caaccccagg acagagctgg agccagggcc 180
 179 agctggatgc ccatgttcca gaggcgaagg aggcgagaca cccacttccc catctgcatt 240
 180 ttctgtcgcg gctgctgtca tcgatcaaag tgtgggatgt gctgcaagac gtagaaccta 300
 181 cctgccctgc ccccgcccc tcccttctct atttattcct gctgccccag aacataggtc 360
 182 ttggaataaa atggctggtt cttttgtttt caaaaaaaaa aaaa 404

184 <210> SEQ ID NO: 6

185 <211> LENGTH: 587

186 <212> TYPE: PRT

187 <213> ORGANISM: Homo sapiens

188 <220> FEATURE:

W--> 189 <221> NAME/KEY:

190 <223> OTHER INFORMATION: 1851405

191 <300> PUBLICATION INFORMATION:

192 <400> SEQUENCE: 6

193 Met Ser Asn Val Met Gly Asn Thr Thr Trp Thr Thr Ser Gly Leu
 194 1 5 10 15
 195 Lys Ser Gln Gly Arg Leu Ser Val Gly Ser Asn Arg Asp Arg Glu
 196 20 25 30
 197 Ile Ser Met Ser Val Gly Leu Gly Arg Ser Gln Leu Asp Ser Lys
 198 35 40 45
 199 Gly Gly Val Val Gly Gly Thr Ile Asp Val Asn Ala Leu Glu Met

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/838,044

DATE: 01/25/2002

TIME: 16:50:53

Input Set : N:\Crf3\RULE60\09838044.raw

Output Set: N:\CRF3\01252002\I838044.raw

200		50		55		60									
201	Val	Ala	His	Ile	Ser	Glu	His	Pro	Asn	Gln	Gln	Pro	Ser	His	Lys
202					65					70					75
203	Ile	Gln	Ile	Thr	Met	Gly	Ser	Thr	Glu	Ala	Arg	Val	Asp	Tyr	Met
204					80					85					90
205	Gly	Ser	Ser	Ile	Leu	Met	Gly	Ile	Phe	Ser	Asn	Ala	Asp	Leu	Lys
206					95					100					105
207	Leu	Gln	Asp	Glu	Trp	Lys	Val	Asn	Leu	Tyr	Asn	Thr	Leu	Asp	Ser
208					110					115					120
209	Ser	Ile	Thr	Asp	Lys	Ser	Glu	Ile	Phe	Val	His	Gly	Asp	Leu	Lys
210					125					130					135
211	Trp	Asp	Ile	Phe	Gln	Val	Met	Ile	Ser	Arg	Ser	Thr	Thr	Pro	Asp
212					140					145					150
213	Leu	Ile	Lys	Ile	Gly	Met	Lys	Leu	Gln	Glu	Phe	Phe	Thr	Gln	Gln
214					155					160					165
215	Phe	Asp	Thr	Ser	Lys	Arg	Ala	Leu	Ser	Thr	Trp	Gly	Pro	Val	Pro
216					170					175					180
217	Tyr	Leu	Pro	Pro	Lys	Thr	Met	Thr	Ser	Asn	Leu	Glu	Lys	Ser	Ser
218					185					190					195
219	Gln	Glu	Gln	Leu	Leu	Asp	Ala	Ala	His	His	Arg	His	Trp	Pro	Gly
220					200					205					210
221	Val	Leu	Lys	Val	Val	Ser	Gly	Cys	His	Ile	Ser	Leu	Phe	Gln	Ile
222					215					220					225
223	Pro	Leu	Pro	Glu	Asp	Gly	Met	Gln	Phe	Gly	Gly	Ser	Met	Ser	Leu
224					230					235					240
225	His	Gly	Asn	His	Met	Thr	Leu	Ala	Cys	Phe	His	Gly	Pro	Asn	Phe
226					245					250					255
227	Arg	Ser	Lys	Ser	Trp	Ala	Leu	Phe	His	Leu	Glu	Glu	Pro	Asn	Ile
228					260					265					270
229	Ala	Phe	Trp	Thr	Glu	Ala	Gln	Lys	Ile	Trp	Glu	Asp	Gly	Ser	Ser
230					275					280					285
231	Asp	His	Ser	Thr	Tyr	Ile	Val	Gln	Thr	Leu	Asp	Phe	His	Leu	Gly
232					290					295					300
233	His	Asn	Thr	Met	Val	Thr	Lys	Pro	Cys	Gly	Ala	Leu	Glu	Ser	Pro
234					305					310					315
235	Met	Ala	Thr	Ile	Thr	Lys	Ile	Thr	Arg	Arg	Arg	His	Glu	Asn	Pro
236					320					325					330
237	Pro	His	Gly	Val	Ala	Ser	Val	Lys	Glu	Trp	Phe	Asn	Tyr	Val	Thr
238					335					340					345
239	Ala	Thr	Arg	Asn	Glu	Glu	Leu	Asn	Leu	Leu	Arg	Asn	Val	Asp	Ala
240					350					355					360
241	Asn	Asn	Thr	Glu	Asn	Ser	Thr	Thr	Val	Lys	Asn	Ser	Ser	Leu	Leu
242					365					370					375
243	Ser	Gly	Phe	Arg	Gly	Gly	Ser	Ser	Tyr	Asn	His	Glu	Thr	Glu	Thr
244					380					385					390
245	Ile	Phe	Ala	Leu	Pro	Arg	Met	Gln	Leu	Asp	Phe	Lys	Ser	Ile	His
246					395					400					405
247	Val	Gln	Glu	Pro	Gln	Glu	Pro	Ser	Leu	Gln	Asp	Ala	Ser	Leu	Lys
248					410					415					420

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/838,044

DATE: 01/25/2002

TIME: 16:50:54

Input Set : N:\Crf3\RULE60\09838044.raw

Output Set: N:\CRF3\01252002\I838044.raw

L:21 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:72 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:2
L:95 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:139 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:146 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:4
L:172 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5
L:189 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:6
L:278 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
L:318 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:8
L:344 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:9
L:348 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:365 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:10
L:383 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:11
L:398 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:12
L:413 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:13
L:430 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:14